DOCUMENT RESUME

ED 314 440 TM 014 233

AUTHOR Schmitt, Dorren Rafael
TITLE Three Methods of Cross-Validation: Is There a Difference in Conclusions?
PUB DATE Nov 89
NOTE 21p.; Paper presented at the Annual Meeting of the Mid-South Educational Research Association (Little Rock, AR, November 8-10, 1989).
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Comparative Analysis; Educational Research; *Generalizability Theory; Methods Research; *Regression (Statistics); *Research Methodology; Sample Size; Validity
IDENTIFIERS *Cross Validation; *Invariance Principle

ABSTRACT

Invariance procedures have been known for over three decades. Through the years, these procedures have not been widely used by researchers. One reason for the lack of use is that most of the articles on invariance procedures have been mathematically oriented. This mathematical orientation and the lack of mathematical background on the part of most educational researchers have inhibited the use of generalizability procedures. Generalizability procedures are important in educational research due to the frequency of researchers needing to use small sample sizes. Problems with the use of small samples are compounded by the frequent use of statistical techniques that capitalize on chance. This paper briefly discusses the history of invariance procedures. A discussion of three methods of cross-validation and illustrations for regression are also presented. The three methods are standard cross-validation, double cross-validation, and Gollob's procedure. These procedures can be performed with univariate as well as multivariate statistical procedures. It is contended that, if the researcher cannot use an independent sample, either standard cross-validation or Gollob's procedure would produce the same results. Unless the sample has less than 30 cases, use of standard cross-validation is recommended. Six data tables are provided. (Author/TJH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED314440

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

✓ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Dorren Rafael Schmitt

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

THREE METHODS OF CROSS-VALIDATION: IS THERE A DIFFERENCE IN CONCLUSIONS?

Dorren Rafael Schmitt

Edcom

Paper presented at the annual meeting of the Mid-South Educational Research Association, Little Rock, AR, November 1989.

TM014233

BEST COPY AVAILABLE

ABSTRACT

Invariance procedures have been known for over three decades. Through the years, these procedures have not been widely employed by researchers. One reason for the lack of use is that the majority of the articles on invariance procedures have been mathematically oriented. The mathematical orientation of articles and the lack of mathematical background on the part of most educational researchers have inhibited the use of generalizability procedures.

Generalizability procedures are important in educational research due to the frequency of researchers needing to use small sample sizes. The use of small samples is compounded by the frequent use of statistical techniques which capitalize on chance.

This paper briefly discusses the history of invariance procedures. A discussion of three methods of cross-validation and illustrations for regression will be presented. Finally, a discussion for what differences in the conclusions have resulted from the three methods.

Three methods of cross-validation: Is there a difference in conclusions?

The generalizability of research results have been under question for the past four decades years. Most discussions have generally been found within the realm of statisticians. The papers have been heavily mathematical based and difficult of the general researching public to read. Thus, cross-validation and other generalizability methods have mostly stayed in the closet. But, these procedures can be useful for researchers.

Generally, a researcher wishes to generalize to the population from which a sample was drawn. But, results are open to skepticism if the sample may not be representative of the population or when the analytic methods employed require larger sample sizes than the research can utilize (Thompson, 1988). In addition, questions can be raised as to the reproducibility of the results because of the analytical methods used. But selected analytic methods can help resolve questions regarding the generalizability of results.

The paper discusses reasons why these generalizability analytic procedures are important. Also, three common methods of the cross-validation generalizability procedure are reviewed. Concrete examples are provided using regression to illustrate the methods are presented.

Sample Representativeness

The first reason to use invariance procedures is to determine bounds for the results. Sample size is an important aspect to be considered for all parametric statistics, and even more so for multivariate statistics. Multiple dependent variables and their reliability increases a larger sample size for multivariate statistics. As variables, both independent and dependent, are added to the equation, there is more question about the reliability of the results. A large sample is needed to off-set the question of reliability.

A small sample may not truly represent the population. Or, the sample may introduce questions about the reliability of the results due only to sample size. Thus, sample size is one design aspect that will determine if invariance procedures are important. Invariance procedures can aid in establishing some confidence and generalizability regarding obtained results.

When a small sample is used, the two invariance procedures recommended are bootstrap and jackknife. Bootstrap, which is a procedures that concatenates the data set, and jackknife, which divides the actual data set into multiple data sets, would be recommended. These procedures are recommended since the size of the data set is not significantly reduced by the methods.

The use of multiple invariance procedures in a given study is also recommended. Thompson (1984) stated that two procedures would provide with a range in with the results would probably fall

within. A confidence level would be created. The use of two procedures is not only helpful when the researcher has a small sample size, but also when correlational statistics (i.e., factor analysis, canonical correlation, Pearson correlation, OVA procedures) are employed (Thompson, 1984). Statistical methods tend to capitalize on chance occurrences and would provide an inflated result (Nunnally, 1967; Thompson, 1984). The capitalization is greatest with the OVA techniques, since data is lost often due to reducing interval data to nominal data. The invariance procedures would establish some bounds to guide interpretation.

Analytic Methods Requirements

Generalizability procedures are needed with some analytic methods more so than others. The multivariate statistical methods tend to capitalize on sampling error even more so than the univariate methods, since these methods are more sensitive to variation (Fornell, 1978). Multivariate statistics tend to be more sensitive to outliers, more so than univariate statistics. In addition to the type of analytic procedure the sample size needed for the analysis leads many researchers to use generalizability procedures. Most specifically, multivariate procedures tend to require larger sample sizes to provide more stability because of increased reliability variance and the sensitivity to variation (Balian, 1982; Katzell, 1951). Researchers may not have the luxury

of a large sample and thus can rely on generalizability procedures to increase confidence in their interpretations.

All the generalizability procedures can be used on all parametric statistics. This is from as simple as the descriptive statistics of means and standard deviations through the more sophisticated multivariate procedures including factor analysis and canonical correlation. For the purposes of simplicity, regression analysis will be used to perform all examples.

Basis for Examples

The data sets is used for the cross-validation examples are of 30 college student questionnaires were used. Student grades point average (gpa) for all remedial and/or freshman English and mathematics courses was used for the dependent variable. The student's high school gpa, usage of the university resource center, and the level of class currently enrolled in were all independent variables. Table 1 presents the two data sets. The first data set will be used for all three cross-validation procedures. The second data set is used for the double cross-validation example since two independent samples are required to that particular procedure.

INSERT TABLE 1 ABOUT HERE

Three Cross-Validation Procedures

The cross-validation procedure has been available to researchers prior to 1951. Mosier (1951) and Katzell (1951) discussed aspects of cross-validation for the general researcher rather than for a statistician. Katzell (1951) argued that cross-validation was not used widely or recognized in the literature. Krus and Fuller (1982) stated that, after many papers, cross-validation is not frequently used partly due to not understanding the fundamentals of the procedure.

For several years, the purpose of cross-validation seemed unclear. Horst (1966) pointed out that cross-validation does not directly assess the generalizability of the best possible equation or factor structure for a particular sample, since replication would be the acid test of generalizability. But cross-validation by randomly splitting the sample into subsets does at least provide some valid assessment of the equation derived (Drehmer & Morris, 1981), a workable method for estimating the degree of shrinkage (Kerlinger & Pedhazur, 1973), and is favored by correlationists (Cronbach, 1957).

There are three methods of cross-validation to determine reproducibility or generalizability include: (a) standard cross-validation, (b) double cross-validation, and (c) Gollob's Procedure. Each of these techniques introduce a different type of sampling logic.

Standard Cross-validation

The **standard cross-validation** procedure calls for the researcher to split the actual data set into random samples. Gorsuch (1983, p. 334) recommends the cross-validation process be used in factor-analytic [or other correlational, such as regression] studies to

> replicate the analysis by randomly splitting the sample in half, factoring both halves and identifying the factors that match across the two samples.

Thus, the particular analytical technique employed for the entire data set is redone on the two random samples. The weights derived from the analysis of the two random samples are crossed or, in the case of a factor analysis, the factors are compared. Synthetic variables are created for the random sample using the actual weights and the crossed weights. The synthetic variables are compared using a Pearson product-moment correlation. If the variables are "highly" correlated, then the sample is said to be invariant. Thus, the results would be generalizable. Otherwise, the results would not be deemed generalizable.

The example for the cross-validation uses the first data set with the analytic procedure regression. First the analysis is performed on the entire data set. The data set is then randomly split in half. The regression procedures is then run separately on the two hales of the original data set. The regression equations for the three runs are:

Total: GPA′ = (HS * -.106) + (CLASS * .432) + (USE * .008) + 2.447

Sub1: GPA′ = (HS * -.151) + (CLASS * -.188) + (USE * .030) + 2.370

Sub2: GPA′ = (HS * 1.036) + (CLASS * .973) + (USE * .004) + 2.012

Next, the weights for the two subgroups are then crossed. Thus, the subsample 1′s regression equation is used to created crossed gpa′ for subsample 2 and the same for subsample 1 with subsample 2′s regression equation. The synthetic variables are then compared using a Pearson correlation. Table 2 presents the synthetic variables created by the crossing of the regression weights. The correlation matrix is presented in Table 3. Those scores which compares the same root with the group and weight cross are highlighted in bold. The synthetic variables were not highly correlated with each other. The synthetic variables for the second group correlated higher than for the first group. Thus, the results are not invariant.

INSERT TABLE 2 AND 3 ABOUT HERE

Double Cross-Validation

The **double cross-validation** method is used when the researcher can collect two independent random samples of equal size. The researcher analyzes each sample independently, determine the weights, and consider the results separately. Then, each set of weights is applied back on the **other** sample to compute the predictive effectiveness.

The main difference between the standard or simple cross-validation and the double cross-validation is the number of independent samples. The simple has only one sample and two smaller samples are drawn for that original sample. On the other hand, double begins with two independent samples and no subsample is ever drawn.

The assets of double cross-validation is that (1) all of the cases are used rather than a portion; (2) fewer sample deviations are created; (3) the procedures is more stable and more valid; and (4) two unbiased estimates are established, each on a larger sample than if the conventional cross-validation method was employed. Thus, the research must weight the need of two independent samples with the four advantages cited above.

To illustrate double cross validation, the two data sets in Table 1 were used. The two data sets provided the following regression equations:

Set1: GPA′ = (HS * -.106) + (CLASS * .432) + (USE * .008) + 2.447

Set2: GPA′ = (HS * .143) + (CLASS * .296) + (USE * .002) + 2.251

As with standard cross-validation, the weights are crossed for the two sets. The synthetic variables created by the crossing of the weights are used to determine if the results are invariant. Table 4 presents the correlation of the synthetic variables.

INSERT TABLE 4 ABOUT HERE

Notice that the correlation between the actual gpa and all of the predicted gpa's are low. But, notice that the correlation between the crossed weights are high. Thus, using double cross-validation, the regression analysis results are considered invariant.

Gollob's Procedure

Gollob (1967) proposed an empirical estimate of population cross-validity for small sample situations. The first step is to calculate the predicted criterion score for each observation of the sample. The prediction for each observation is accomplished by removing the observation from the sample and running a regression analysis on the remaining observations, N-1 (Drehmer & Morris, 1981). The equation is used to estimate the criterion score for the specific observation omitted from the analysis.

The correlation between the estimated and actual criterion values are computer after all estimated criterion scores have been calculated. The correlation is the estimated population cross-validity of the regression equation.

The major drawback to Gollob's procedure is the inordinate amount of computations that must be performed by hand. Much of the analysis must be done by hand. To relieve some of the computations Drehmer and Morris (1981) developed a FORTRAN program for performing the analysis for regression. The following example and discussion was done without the use of that program.

Although Drehmer and Morris (1981) suggest that Gollob's

procedure is designed for small samples of 20 to 60 cases with 3 to 5 predictors, this researcher would not like to perform this procedure with more than the 30 subjects and 3 predictors involved in this example. The computational time is about 3 to 5 times longer than the previous two examples.

The first data set was used for the Gollob Procedure. Thirty-one regression procedures were run. The first procedure used all 30 individuals. The remaining 30 procedures used 29 individual. The regression procedure that omitted the given case, such as case 1, would be the equation to predict the criterion variable for that omitted case, such as for case 1. Table 5 presents the 30 regression equations and the calculations for the predicted score for each case.

INSERT TABLE 5 ABOUT HERE

Following the calculations of the predicted scores, the actual score was correlated with the predicted score. If there is a high correlation, then the results are said to be invariant or generalizable. Table 6 presents the correlation of the actual and synthetic variables. Since the correlation is low (-.175) the results are said not to be invariant or not generalizable.

INSERT TABLE 6 ABOUT HERE

Summary

As discussed above, there are different sampling procedures for cross-validation. The three methods discussed in this paper were standard and double cross-validation and Gollob's procedure. There are other methods and variations on these discussed. These procedures can be performed for the univariate as well as the multivariate statistical procedures.

Differing results were found. The standard cross-validation and Gollob's procedure found that the results were not invariant. Both of these procedures used the same sample. The one procedure that used two samples was the double cross-validation procedure. The double cross-validation determined that the results were generalizable. Thus, there seems to be some differences if the researcher estimates using one sample or usee two independent samples.

So it seems that if the research can not use an independent sample that either standard cross-validation or Gollob's procedure would produce the same results. Unless the sample has fewer than 30 cases, this researcher would recommend using standard cross-validation described in this paper or some other variation of it. The computational effort of Gollob's procedure is excessive for the same conclusions reached by standard cross-validation.

References

Balian, E. S. (1982). How to design, analyze, and write doctoral research. New York: University Press of America.

Cronbach, L. J. (1957). The two disciplines of scientific psychology. American Psychologist, 12, 671-684.

Drehmer, D. E., & Morris, G. W. (1981). Cross-validation with small samples: An algorithm for computing Gollob's estimator. Educational and Psychological Measurement, 41, 195-200.

Fornell, C. (1978). Three approaches to canonical analysis. Journal of Market Research Society, 20, 166-181.

Horst, P. (1966). An overview of the essentials of multivariate analysis method. In R. B. Cattell, (Ed.) Handbook of multivariate experimental psychology. Chicago: Rand-McNally.

Katzell, R. A. (1951). Symposium III: Cross-validation of item analysis. Educational and Psychological Measurement, 11, 16-22.

Kerlinger, F. M., & Pedhazur, E. J. (1973). Multiple regression in behavioral research. New York: Holt.

Krus, D. J., & Fuller, E. A. (1982). Computer assisted multicrossvalidation in regression analysis. Educational and Psychological Measurement, 42, 187-193.

Mosier, C. I. (1951). Symposium I: Problems and designs of cross-validation. Educational and Psychological Measurement, 11, 5-11.

Nunnally, J. C. (1967). Psychometric theory. New York: McGraw-Hill.

Thompson, B. (1984). Canonical correlation analysis: Uses and interpretation. Beverly Hills: SAGE.

Table 1

Data Sets
Used of Cross-Validation

Independent Sample 1					Independent Sample 2			
GPA	H.S.	CLS	USE	GP2	GPA	H.S.	CLS	USE
2.0	3.3	1	30	1	2.2	2.1	1	12
3.5	2.9	1	10	2	2.5	2.4	0	13
1.7	2.5	0	09	2	2.7	2.5	0	19
2.2	2.0	0	15	1	2.2	2.0	0	10
1.2	2.1	0	27	2	3.2	2.1	0	07
2.7	3.5	0	50	1	2.7	3.0	0	50
4.0	3.7	1	36	2	4.0	3.1	1	21
2.1	2.0	0	23	1	3.1	3.0	0	27
3.3	2.5	1	42	2	3.7	2.9	1	22
2.9	2.2	0	14	1	2.3	3.2	0	11
2.4	3.4	0	69	2	2.6	3.6	0	41
2.8	3.8	0	17	1	3.8	3.0	0	12
4.0	3.2	1	37	1	3.0	2.2	1	17
3.6	2.9	1	32	1	3.3	2.1	1	34
2.9	2.9	1	27	2	2.6	2.3	1	29
2.3	3.0	0	03	2	3.3	2.4	0	73
2.5	3.7	1	07	2	3.5	3.2	1	02
2.7	2.6	1	12	1	2.2	3.6	1	42
3.8	2.1	1	05	2	3.5	2.6	1	15
2.2	2.0	0	17	1	2.0	2.3	0	19
4.0	2.5	1	03	2	4.0	3.5	1	33
1.1	3.6	1	08	1	3.4	4.0	1	19
2.6	2.3	0	33	2	2.1	2.8	0	37
2.1	2.1	1	21	1	2.1	3.1	1	51
3.3	2.5	1	11	2	3.6	2.1	1	31
0.7	3.2	1	23	2	2.7	2.2	1	03
3.1	2.9	0	15	1	2.1	1.9	0	15
2.4	2.1	0	02	1	2.9	2.7	0	22
1.6	2.2	1	09	2	2.6	3.2	1	19
1.3	3.4	1	19	1	1.5	3.0	1	04

GPA -- GRADE GIVEN FOR MATHEMATICS CLASS
CLS -- TYPE OF CLASS: REMEDIAL 0 COLLEGE LEVEL 1
USE -- USAGE OF THE LRC SERVICES
GRP -- SUBGROUP FOR THE STANDARD CROSS-VALIDATION

Table 2
GPA and Synthetic Variables
Created from the Crossing of Regression Weights

CASE	GPA	GPA11	GPA12	GPA21	GPA22
1	2.000	2.584	2.986	.	.
2	3.500	.	.	2.044	2.921
3	1.700	.	.	2.263	1.958
4	2.200	2.518	2.000	.	.
5	1.200	.	.	2.863	2.044
6	2.700	3.342	2.086	.	.
7	4.000	.	.	2.703	2.996
8	2.100	2.758	2.032	.	.
9	3.300	.	.	3.065	3.063
10	2.900	2.458	1.989	.	.
11	2.400	.	.	3.927	2.166
12	2.800	2.306	1.943	.	.
13	4.000	2.809	3.018	.	.
14	3.600	2.704	3.009	.	.
15	2.900	.	.	2.554	2.989
16	2.300	.	.	2.007	1.916
17	2.500	.	.	1.833	2.880
18	2.700	2.149	2.939	.	.
19	3.800	.	.	2.015	2.929
20	2.200	2.578	2.008	.	.
21	4.000	.	.	1.895	2.907
22	1.100	1.878	2.887	.	.
23	2.600	.	.	3.013	2.061
24	2.100	2.495	2.993	.	.
25	3.300	.	.	2.135	2.939
26	0.700	.	.	2.389	2.962
27	3.100	2.382	1.968	.	.
28	2.400	2.113	1.944	.	.
29	1.600	.	.	2.120	2.942
30	1.300	2.239	2.939	.	.

Note.
GPA -- THE ACTUAL COLLEGE GPA
GPA11 -- GPA FROM SAMPLE 1'S SCORES AND OWN BETA WEIGHTS
GPA12 -- GPA FROM SAMPLE 1'S SCORES AND SAMPLE 2'S BETA WEIGHTS
GPA21 -- GPA FROM SAMPLE 2'S SCORES AND SAMPLE 1'S BETA WEIGHTS
GPA22 -- GPA FROM SAMPLE 2'S SCORES AND OWN BETA WEIGHTS

Table 3
Pearson Correlation Comparing GPA with the Synthetic Scores
From the Standard Cross-Validation

	GPA	GPA11	GPA12	GPA21	GPA22
GPA	1.000				
GPA11	0.449	1.000			
GPA12	-.061	-.133	1.000		
GPA21	-.138	.	.	1.000	
GPA22	0.443	.	.	-.324	1.000

Note. The "." represents incalculatable correlations since there were no cases in common.
NUMBER OF OBSERVATIONS = 15

Table 4
Pearson Correlation of Actual scores with synthetic scores
From the Double Cross-Validation Analysis

	GPA	GPA11	GPA21	GPA21	GPA22
GPA	1.000				
GPA11	0.242	1.000			
GPA12	0.192	**0.885**	1.000		
GPA21	0.209	.	.	1.000	
GPA22	0.278	.	.	**0.755**	1.000

Note.
GPA -- THE ACTUAL COLLEGE GPA (N=60)
GPA11 -- GPA FROM SAMPLE 1'S SCORES AND OWN BETA WEIGHTS
GPA12 -- GPA FROM SAMPLE 1'S SCORES AND SAMPLE 2'S BETA WEIGHTS
GPA21 -- GPA FROM SAMPLE 2'S SCORES AND SAMPLE 1'S BETA WEIGHTS
GPA22 -- GPA FROM SAMPLE 2'S SCORES AND OWN BETA WEIGHTS
NUMBER OF OBSERVATIONS = 30

Table 5
Computation of estimated scores for Gollob's Procedure

CASE	GPA	GPA′ =	(H.S. * weight)	+ (CLASS * weight)	+ (USE * weight)	+ CONSTANT
1	2.0	2.841	(3.3 * -.080)	+ (1.0 * .477)	+ (30 * .009)	+ 2.358
2	3.5	2.605	(2.9 * -.119)	+ (1.0 * .387)	+ (10 * .010)	+ 2.449
3	1.7	2.323	(2.5 * -.100)	+ (0 * .379)	+ (9 * .007)	+ 2.510
4	2.2	2.377	(2.0 * -.116)	+ (0 * .422)	+ (15 * .008)	+ 2.489
5	1.2	2.601	(2.1 * -.198)	+ (0 * .358)	+ (27 * .010)	+ 2.747
6	2.7	2.789	(3.5 * -.124)	+ (0 * .457)	+ (50 * .007)	+ 2.493
7	4.0	2.585	(3.7 * -.203)	+ (1.0 * .371)	+ (36 * .006)	+ 2.749
8	2.1	2.454	(2.0 * -.132)	+ (0 * .371)	+ (23 * .008)	+ 2.534
9	3.3	2.865	(2.5 * -.071)	+ (1.0 * .391)	+ (42 * .006)	+ 2.399
10	2.9	2.270	(2.2 * -.085)	+ (0 * .476)	+ (14 * .009)	+ 2.331
11	2.4	2.824	(3.4 * -.097)	+ (0 * .407)	+ (69 * .011)	+ 2.395
12	2.8	1.954	(3.8 * -.233)	+ (0 * .543)	+ (17 * .010)	+ 2.658
13	4.0	2.708	(3.2 * -.117)	+ (1.0 * .349)	+ (37 * .005)	+ 2.548
14	3.6	2.749	(2.9 * -.090)	+ (1.0 * .373)	+ (32 * .006)	+ 2.445
15	2.9	2.783	(2.9 * -.105)	+ (1.0 * .424)	+ (27 * .008)	+ 2.447
16	2.3	2.129	(3.0 * -.120)	+ (0 * .452)	+ (3 * .009)	+ 2.453
17	2.5	2.554	(3.7 * -.100)	+ (1.0 * .433)	+ (7 * .008)	+ 2.435
18	2.7	2.699	(2.6 * -.106)	+ (1.0 * .432)	+ (12 * .008)	+ 2.447
19	3.8	2.526	(2.1 * -.015)	+ (1.0 * .332)	+ (5 * .010)	+ 2.175
20	2.2	2.393	(2.0 * -.118)	+ (0 * .421)	+ (17 * .008)	+ 2.493
21	4.0	2.482	(2.5 * -.073)	+ (1.0 * .343)	+ (3 * .011)	+ 2.288
22	1.1	2.848	(3.6 * .068)	+ (1.0 * .467)	+ (8 * .003)	+ 2.112
23	2.6	2.462	(2.3 * -.098)	+ (0 * .440)	+ (33 * .008)	+ 2.423
24	2.1	2.936	(2.1 * -.187)	+ (1.0 * .510)	+ (21 * .009)	+ 2.630
25	3.3	2.656	(2.5 * -.084)	+ (1.0 * .389)	+ (11 * .009)	+ 2.378
26	0.7	2.866	(3.2 * -.041)	+ (1.0 * .541)	+ (23 * .008)	+ 2.272
27	3.1	2.168	(2.9 * -.150)	+ (0 * .522)	+ (15 * .010)	+ 2.453
28	2.4	2.214	(2.1 * -.101)	+ (0 * .447)	+ (2 * .009)	+ 2.408
29	1.6	2.859	(2.2 * -.189)	+ (1.0 * .529)	+ (9 * .007)	+ 2.683
30	1.3	2.812	(3.4 * -.016)	+ (1.0 * .490)	+ (19 * .007)	+ 2.243

GPA′ -- the estimated GPA score

Table 6
Pearson Correlation of GPA with Estimated GPA
by Gollob's Procedure

	GPA	GPA′
GPA	1.00	
GPA′	-.175	1.00